Ming Hou*
Justin G. Hollands
Defence Research & Development Canada—Toronto

Andrea Scipione
CAE Professional Services Canada

Lochlan Magee
Defence Research & Development Canada—Toronto

Mike Greenley
CAE Professional Services—Canada

Comparative Evaluation of Display Technologies for Collaborative Design Review

Abstract

The effectiveness of five display technologies for supporting a collaborative workspace design review was compared. Participants searched for design flaws in a model of the front dashboard of a vehicle including an in-vehicle navigation system. The display types were 2D CRT, 3D CRT, 3D via Curved plasma display, a large DataWall display, and a cave automatic virtual environment (CAVE). Detection accuracy, time, and usability measures were obtained. The results indicated that detection accuracy was higher for 3D CRT and Curved displays than the 2D display or more immersive DataWall and CAVE displays. Additionally, a speed-accuracy trade-off was observed such that detection time was longer for 3D CRT and Curved displays than for 2D, or the more immersive displays. Subjective measures revealed that participants' comfort and confidence level was lower with the 2D displays than the 3D displays. Lack of sufficient training time is likely to have affected detection accuracy with the more immersive 3D displays. Overall, the use of the 3D CAD model on a standard CRT or a Curved display was the most cost-effective for collaborative design review.

1 Introduction

1.1 Background

In developing a product or system, a prototype is often reviewed collaboratively by a group of users in a design-test-redesign cycle. Collaborative design reviews can improve the design process by enabling broad expertise to influence the product or system evaluation. Immersive display technologies offer a collaborative design team a means to present and interact with a design prototype that has the potential to improve the effectiveness of the review process. Display technologies are available in varying degrees of fidelity, and a prototype can be presented in the form of a two-dimensional (2D) blueprint, a three-dimensional (3D) image, or an immersive virtual environment (VE). Compared to a 2D presentation, a computer-generated VE can offer a greater sense of presence or immersion. Immersion can be defined as the extent to which users are convinced that they are somewhere other than where they physically are (Wickens & Hollands, 2000). An immersive VE could preserve

Presence, Vol. 18, No. 2, April 2009, 125–138

*Correspondence to Ming.Hou@drdc-rddc.gc.ca.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **APR 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Comparative Evaluation of Display Technologies for Collaborative Design Review**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defence Research & Development Canada?Toronto,1133 Sheppard Avenue West,P.O. Box 2000,Toronto, Ontario, Canada M3M 3B9,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**14**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

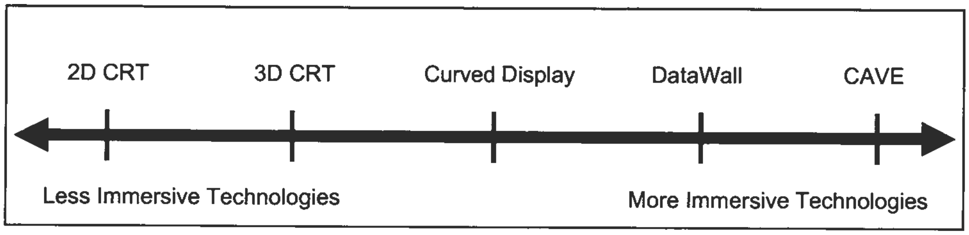


Figure 1. *A continuum of immersive display technologies.*

many characteristics of a product or a system so that the environment represents a realistic model of the product. In addition, computer-generated artificial VEs can be interactively manipulated and experienced, which may make it easier to detect costly design flaws.

Current VE displays vary along a continuum of immersion, as illustrated in Figure 1. Toward the left end of the continuum are less immersive technologies. Here, a 3D object is viewed from a single viewpoint or a set of discrete 2D viewpoints using a standard flat monitor, such as a cathode ray tube (CRT) screen. Interactions with such technologies typically involve conventional input devices, such as keyboards and mice. Some less immersive display technologies give users a wide field of view, real-time interaction, and a sense of presence, but not total immersion. These include 3D representations on flat CRTs and some Curved display systems with more compelling depictions of space and depth than flat displays. Interaction with such displays can be provided by different input devices such as a mouse, trackball, and joystick. Such less immersive technologies do not require a high level of graphics or special hardware, making them a low-cost option. Toward the right end of the continuum in Figure 1 are more immersive technologies that use a variety of environmental features to provide users with the illusion of immersion. These include 3D viewing, real-time dynamics, an ego-centered frame of reference, multimodal interaction, and a wide field of view (Wickens & Hollands, 2000). Interactions with immersive technologies can be provided by a variety of input devices, such as a joystick or Pinch glove. An example is the large, flat DataWall projector that users can interact with by using projected stereo images with wide fields of view. Another example is the cave automatic virtual environment (CAVE), where users can move freely and interact with stereo images projected on four to six walls of a room with unlimited field of view (Cruz-Neira, Sandin, & DeFanti, 1993; Dynott et al., 2002; Owen, 1999).

Immersive display technologies show great potential for collaborative design review. First, a 3D display representation preserves characteristics of a real-world object more realistically than a 2D representation; thus, it provides a more realistic view than can be obtained through a set of 2D elevations (Wickens & Hollands, 2000). Second, an ego-centered and immersed viewpoint provides a more "natural" feeling than a 2D representation when interacting with a display representation (McCormick, Wickens, Banks, & Yeh, 1998; Peterson, Wells, Furness, & Hunt, 1998). Third, more immersive display technology provides users with multimodal interaction including auditory, kinesthetic, force, and tactile feedback using a Pinch glove, joystick, or other input devices (e.g., Dede, Salzman, & Loftin, 1996; Werkhoven & Groen, 1998). Fourth, some more immersive display technologies provide stereo imagery, allowing users to take advantage of stereopsis to develop a more precise spatial understanding or visualization (Barfield, Zeltzer, Sheridan, & Slater, 1995; Bowman et al., 2002, Snow & Williges, 1997). Fifth, some immersive display technologies track and update the visual scene based on an observer's head or eye movements (Kocian & Task, 1995). These features produce a better sense of presence in an immersive display environment than can be provided by a 2D display. Therefore, when navigating through, interacting with, or understanding 3D environments, users should benefit from more immersive 3D displays.

However, 3D displays can impose certain costs, and some task requirements neutralize their benefits. Additionally, 3D displays may degrade performance, even for tasks that people need to perform in 3D environments (Wickens, 2000). For example, when more immersive display technologies are chosen, there is a cost associated with the keyhole effect, where viewers see only a portion of the total environment at a time (Woods, 1984). When the full environment is represented in a single display, a searching task can be efficiently accomplished exclusively via eye movements, while involving minimal cognitive or physical effort. In some systems, a search must be accomplished using a control interface, as in panning the field of view of a VE system. Greater cognitive and physical effort involved can lead to an incomplete search, and/or an overemphasis on locating objects within the initial forward view. Costs more to immersive display technologies have been found in aviation (Olmos, Wickens, & Chudy, 2000), in battlefield displays (Wickens, Thomas, Merlo, & Hah, 1999), and in data visualization (McCormick et al., 1998). Selecting a display for a design review should consider the benefits and costs of display technologies that vary in degree of immersion, in order to address one of the most critical challenges in applying VE technologies: "choosing which display best fits each application" (Brooks, 1999, p. 27).

1.2 Research Motivation

During the past decade, there has been a great deal of effort in the VE community aimed at developing new displays and improving existing display types. However, there is little work that objectively compares team behavior and performance in different VE displays. Bolstad and Endsley (1999) examined the use of shared mental models among air commanders using shared 2D workstation displays as a means of enhancing team situation awareness. They found that shared displays for a two-person team helped to build shared mental models, but hindered performance due to the extra time required to process the information in initial trials. However, in later trials, team performance was better for individuals who did not first use shared displays. It is not clear whether this result is applicable to other operational tasks, or how it might be influenced by more immersive displays. A more extensive examination of the various influences of VE technologies on team performance is required. Such research could assist those planning a design review before display technologies are acquired and implemented.

To address these questions, we conducted an experiment with a two-person team performing a collaborative design review on a model of a vehicle interior, including a dashboard and an in-dash navigation system. We compared five different display technologies: 2D CAD model on a desktop CRT, 3D CAD model on a desktop CRT, 3D CAD model on a Curved plasma display, 3D CAD model on a DataWall projector, and 3D CAD model on a CAVE. Because many factors vary together with changes in display type, such as display size, resolution, input device, interaction mode, and so forth, it was not technically possible to keep these factors equivalent. However, the results should indicate the relative performance advantages of commercially available display technologies for collaborative design review.

While acknowledging the limitations of 3D immersive formats, given that the design review task requires 3D spatial understanding, we predicted that team performance (and usability ratings) should improve with increased immersion. This general prediction led to two specific hypotheses. First, 3D displays should be superior to 2D displays because 3D environments can provide greater presence with real-time dynamics and an ego-centered frame of reference. Second, more immersive 3D displays should be superior to less immersive 3D displays because of the wide field of view and multimodal interaction in more immersive environments.

2 Method

2.1 Participants

Eighty participants (42 females and 38 males), aged 18 to 58 years ($M = 23.9$, $SD = 4.24$), were recruited. All participants had normal or corrected-to-normal vision and at least two years of driving experience. None knew about the design or the goals of the

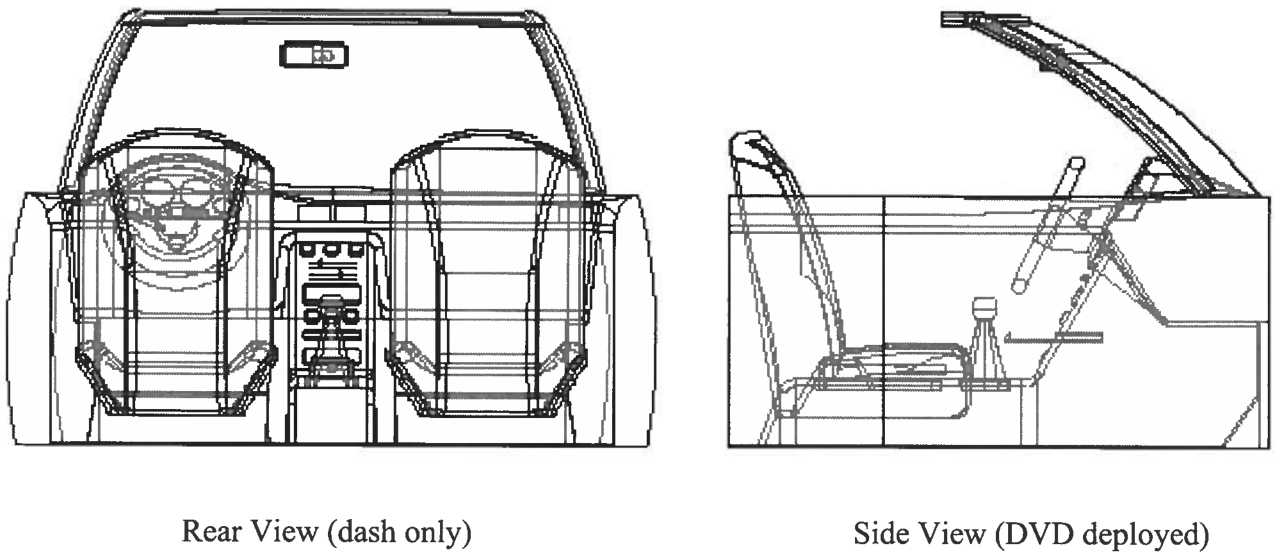


Figure 2. *Example of the 2D CAD model used in the experiment as it appeared on the CRT.*

experiment, or had experience with vehicle manufacturing, or in-vehicle navigation systems. These individuals were randomly paired to produce a set of 40 two-person teams. Each team was randomly assigned to one display technology. There were eight teams (or 16 participants) per display technology.

2.2 Apparatus and Stimuli

The dashboard model included a Digital Video Disc (DVD) player and Liquid Crystal Display (LCD) screen, among other typical dashboard components. Three separate model states were used for exhibiting the navigation system: dash only (without DVD or LCD), DVD tray deployed (LCD not displayed), and LCD in upward position (DVD not displayed). These model states were optionally available for participants for the design review in each condition. Figures 2 and 3 show examples. The 2D CAD model was developed using Pro/ENGINEER (Tickoo, 2002), and converted into PDF files for viewing on a Samsung SyncMaster flat screen MagicBright 17 in CRT monitor. The 2D CAD model included four separate views of the vehicle: behind the vehicle looking forward, above the vehicle looking downward, and from the left and right sides of the vehicle. These four views were available for each of the three model states. The 2D CAD model was static and monochrome, as illustrated in Figure 2. The 2D model was translated into a 3D CAD model using eDrawings (Planchard & Planchard, 2001). The 3D model was shown in dark gray on a blue background. Figure 3 shows examples of the three model states with the 3D CAD model. The same desktop CRT monitor was used in 2D and 3D CRT displays.

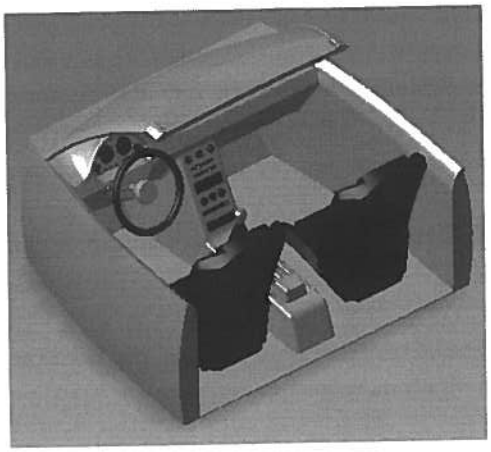
Figure 4 shows examples of the 3D CAD model as viewed using the three other experimental displays: a Curved (hemispheric) Elumens VisionStation plasma display, a Marquis 8500 plasma DataWall projector, and a Fakespace CAVE. The Curved display had a radius of 84 cm (33 in) with a 180° field of view. It allowed participants to see above, below, and to the left and right to produce a realistic sense of space and depth. The Curved display also had a canopy that was closed above the participants during the experiment. The DataWall had a 2,400 cm (96 in) by 1800 cm (72 in) glass screen with 3D CAD images being rear projected. The DataWall projector incorporated a wide field of view while maintaining high resolution. The CAVE was a 303 cm

(120 in) by 303 cm (120 in) by 303 cm (120 in) room with four projected surfaces: front wall, left wall, right wall, and floor. Each projected image was 303 cm (120 in) by 303 cm (120 in) and was rear projected except for the floor which displayed an image projected from above. The 3D CAD models for these three displays were generated on an SGI Onyx 3200 using Open GL Performer. In the DataWall and CAVE environments, the CAD models were stereoscopic and the participants wore StereoGraphics liquid crystal glasses. The characteristics of the various display technologies are summarized in Table 1.

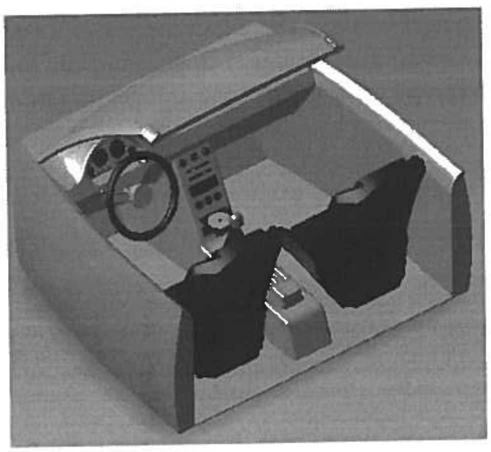
2.3 Control of the CAD Models

A mouse was used to manipulate model orientation and eye reference point for the 3D CRT monitor, the Curved display, and the DataWall projector. In the 2D CRT and all 3D displays (except the CAVE), participants could switch between the three model states (dash only, DVD deployed, and LCD deployed), by clicking the middle mouse button. In the 3D displays (except the CAVE) the models could be translated and rotated in x and y planes with pitch and yaw manipulations using the left and right mouse buttons. When the left mouse button was pressed, the rotation of the models along the x and y axes was controlled by moving the mouse left and right or forward and backward, respectively. Translation along the z-axis (i.e., zooming in and out) was also possible, by pressing the right mouse button and moving the mouse forward and backward. The 2D CAD model did not allow rotation or translation.

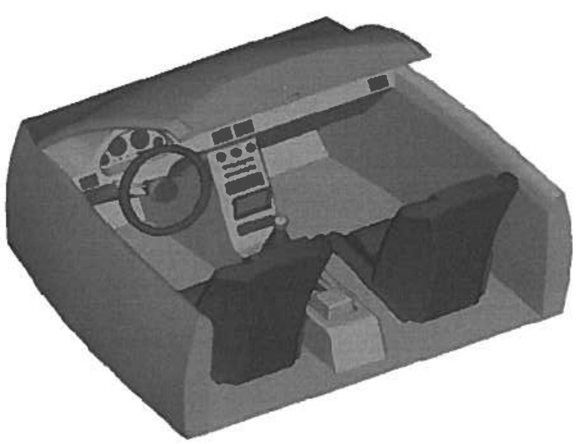
In the CAVE, the model was rendered on the projectors according to the eye point of one (headtracked) participant using a 6 degrees of freedom (DOF) head tracker (Flock of Birds). The other 6 DOF Flock of Birds sensor was attached to a Fakespace Pinch glove which controlled the motion of a virtual hand in the display. The virtual hand mimicked the motion of the



Navigation system flush with dashboard



DVD deployed



Screen deployed in upward position

Figure 3. *Examples of the 3D CAD model used in the experiment as it appeared on the CRT.*

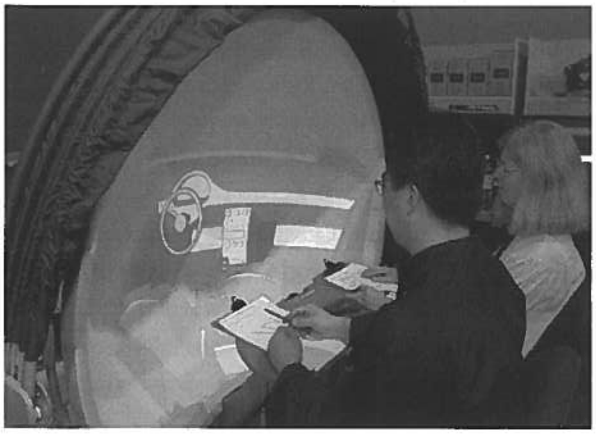
physical hand, and fingertip contacts triggered the switch between the three model states. The model could be translated in and out by moving the glove forward and backward. The model could be rotated in pitch and yaw by rotating the glove along the x and y axes.

The two participants on each team sat side by side during the review process for all display conditions except the CAVE, where they stood side by side. One of the participants controlled the models and the other was responsible for data recording.

2.4 Design and Procedure

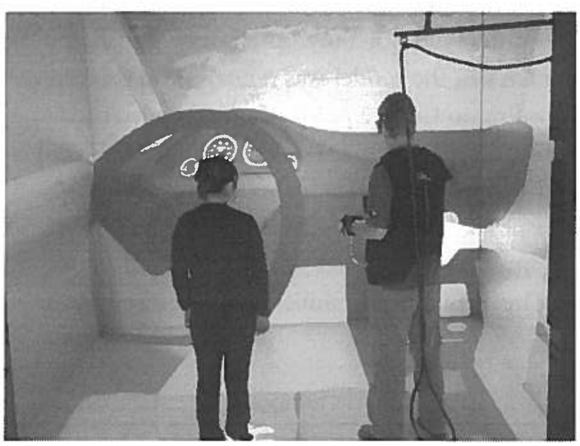
Eight review teams reviewed the CAD model for each display technology. The experiment had a one-way, between-subjects design with five experimental displays. The review task was to search for design flaws embedded in the model that contradicted published recommendations for navigation system installation (Commission of the European Communities, 2000; Green, Levison, Paelke, & Serafin, 1994; Japan Automobile Manufacturers Association, 2000) when using the modeled in-vehicle navigation system. The design flaws, listed in Table 2, were detectable in both 2D and 3D models.

After the screening process, each review team was briefed on the in-vehicle navigation system. Each participant read a brief description of the experiment, the protocol to be followed, and signed an informed consent form. General questions about the experimental design were answered. Each team had about 10 min of familiarization time with the display media and the input device before they started a trial. Participants had a maximum of 20 min to identify as many design flaws in the CAD model as they could. They were encouraged to discuss possible flaws with each other. The trial was



a Curved Elumens VisionStation plasma display

a large DataWall display



a CAVE display

Figure 4. *Examples of 3D CAD models on a Curved plasma display, a large DataWall display, and a CAVE. The canopy is open in the Curved display photograph to provide a view of the display, although it was closed in the experimental trials.*

Table 1. *Characteristics of Display Media for Collaborative Design Review*

Display characteristics	2D CAD on CRT	3D CAD on CRT	Curved plasma display	DataWall projector	CAVE
	DISPLAY MEDIA				
Level of immersion	Less immersive 2D	Less immersive 3D		More immersive 3D	
Display size	17 in	17 in	62.5 in (W) × 57 in (H) × 21 in (D)	96 in × 72 in	4 × 120 in × 120 in
Resolution (pixels)	1280 × 1024 @ 75 Hz refresh rate	1280 × 1024 @ 75 Hz refresh rate	1024 × 768 @ 96 Hz refresh rate	1024 × 768 @ 96 Hz stereo	1024 × 1024 @ 96 Hz stereo
Distance to front display	55 cm sitting	55 cm sitting	55 cm sitting	150 cm sitting	150 cm standing
Physical field of view (approximate)	34 × 26°	34 × 26°	110 × 110°	77 × 62°	210 × 135°
Aspect ratio	4:3	4:3	~5:4	4:3	1:1
Input device	Optical mouse	Optical mouse	SGI mouse	SGI mouse	Fakespace Pinch glove
3D glasses	No	No	No	Yes	Yes
Head tracker	No	No	No	No	Yes 6 DOF Flock of Birds
Model manipulation	Static	Translation, rotation	Translation, rotation	Translation, rotation	Translation, rotation
Display cost (US $ in 2003)	<$1,000	<$1,000	~$29,000	~$200,000	$350,000–$450,000

Table 2. *Design Flaws in In-Car Navigational System*

No.	Flaws
1	System obstructed other vehicle controls and displays
2	Visual display too low (not positioned close to driver's normal line of sight)
3	Controls not in easy reach
4	Frequently used controls not in dominant hand position
5	Safety: location of navigation system distracted driver's attention
6	Poor angle of the screen

completed when the team indicated that they could not identify another design flaw. Time was recorded using a stopwatch. Informal observations of participants' behavior were also noted. After the experiment, participants were debriefed and completed a questionnaire. The experiment, including practice, briefing, and debriefing, lasted approximately 45 min per team.

2.5 Performance Measures

During the experiment, two performance measures were taken: accuracy and time on task. Accuracy was defined by the number of flaws detected. Time on

Table 3. *Questionnaires for Usability Ratings*

	Statements
Review media	It was easy to understand the model/environment. It was easy to identify vehicle features. It was easy to imagine myself interacting with the in-vehicle navigation system. The model and review medium provided a meaningful representation of real-world objects. I was able to visualize the model with correct scaling. I became lost or disorientated when viewing the model.
Review task	It was easy to perform the design review. I felt comfortable when performing the design review. I felt confident in conducting the design review. It was easy to visualize and detect design flaws. I felt confident in the number of design flaws detected.
Quality of collaborative review	I had a strong sense of presence within the car environment. I had a strong sense of "co-presence" with my partner within the car environment. I feel that I worked with my partner to complete the review. I feel that the review medium facilitated and supported collaboration with my partner. I feel that the review medium facilitated communication with my partner.

task was defined by the total time taken after the trial started until no more design flaws could be identified.

2.6 Subjective Measures

Three subjective measures were obtained from the questionnaires: review media usability, review task usability, and quality of collaborative review usability. Six *review media* usability statements examined a review medium's effect on the ease of understanding the vehicle model, identifying features, understanding interaction with the navigation system, and becoming disoriented within the environment. A further statement assessed whether the medium meaningfully represented modeled objects with proper vehicle scaling. Five *review task* usability statements evaluated the ease of performing design review and searching for design flaws, and also how comfortable and confident participants were in conducting the review task and detecting the design flaws. Five *quality of collaborative review* usability statements investigated the perceived level of presence and co-presence within the display environment. The questionnaire also assessed the degree to which a review medium facilitated collaboration and communication between team members. All usability statements are listed in Table 3. Subjective ratings used a 7-point Likert scale, ranging from 1 (Strongly Disagree) to 7 (Strongly Agree). The questionnaire also contained a section for written comments.

3 Results

A one-way between-subjects analysis of variance (ANOVA) was performed on each measure (Howell, 1999). To assess predictions, we used a set of four planned comparisons following the modified Bonferroni

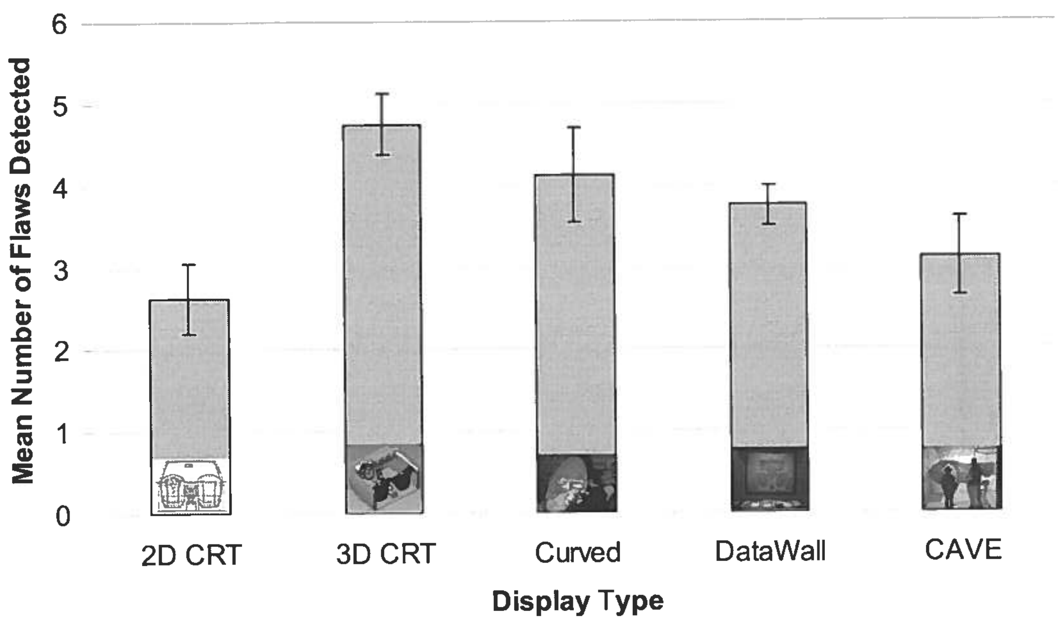


Figure 5. *Accuracy as a function of display type. Error bars indicate standard error for the mean in all graphs.*

procedure (Keppel, 1982, pp. 148–149) which allowed us to use a single alpha level of .05 for all comparisons. The first of these compared the 2D display to all 3D displays combined. The other comparisons grouped the two more immersive 3D displays that included 3D eyewear and specific input devices (DataWall and CAVE), and the two less immersive displays (3D CRT and Curved displays), to keep the number of comparisons manageable. We compared the 2D display to the less immersive 3D displays, and to the more immersive 3D displays. For the final comparison, we compared the less immersive and more immersive 3D displays to each other.

3.1 Performance Measures

3.1.1 Accuracy: Number of Flaws Detected. Figure 5 shows the number of flaws detected in each experimental display. The ANOVA showed that display type affected the number of flaws detected, $F(4, 35) = 3.68, p < .05$. Fewer flaws were detected with the 2D display than with all 3D displays combined (2.6 vs. 3.9), $F(1, 35) = 7.33, p < .05$. Fewer flaws were detected with the 2D display than with the less immersive 3D displays (2.6 vs. 4.4), $F(1, 21) = 11.64, p < .05$. However, there was no significant difference between the 2D display and the more immersive 3D displays (2.6 vs. 3.4), $p > .05$. Finally, more flaws were detected with the less immersive 3D displays than the more immersive 3D displays, (4.4 vs. 3.4), $F(1, 28) = 5.32, p < .05$.

3.1.2 Time on Task: Total Time Spent on Detection Task. Figure 6 shows total time spent on the detection task in each experimental display. Display type affected detection time, $F(4, 35) = 3.98, p < .05$. There was no significant difference in time spent between the 2D display and the 3D displays combined (11.1 vs. 13.7), $p > .05$. However, less time was spent with the 2D display than with the less immersive 3D displays (11.1 vs. 16.8), $F(1, 21) = 7.99, p < .05$. There was no significant difference in time spent between the 2D display and the more immersive displays (11.1 vs. 10.6), $p > .05$. A comparison of the less immersive and more immersive 3D displays showed that less time was spent with the more immersive 3D displays (16.8 vs. 10.6), $F(1, 28) = 13.88, p < .05$.

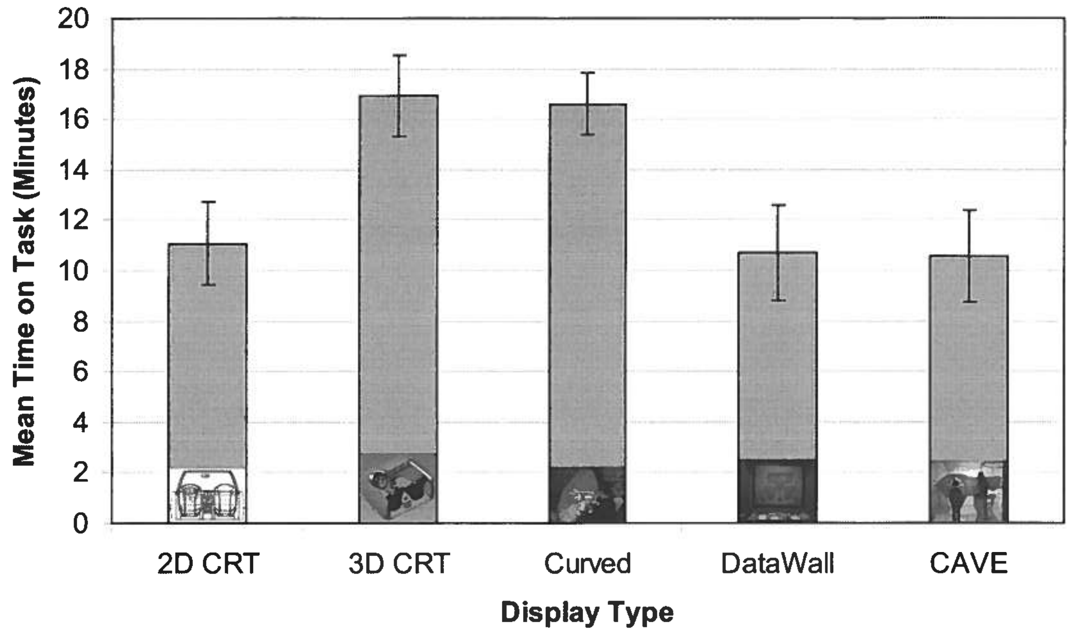


Figure 6. *Time on task as a function of display type.*

3.2 Subjective Measures

Subjective ratings were averaged across the questionnaire statements in Table 3 prior to being submitted to ANOVA. In other words, a usability indication was represented by an overall usability average for each display that was computed from the ratings across the individual usability statements in Table 3. The results are shown in Figure 7.

3.2.1 Review Media Usability. Figure 7 shows the mean review media usability ratings for each experimental display. The ANOVA showed that display type influenced review media usability ratings, $F(4, 35) = 10.08, p < .05$. Participants produced lower review media usability ratings with the 2D display than all 3D displays combined (5.1 vs. 5.9), $F(1, 35) = 4.38, p < .05$. Participants produced lower review media usability ratings with the 2D display than with the less immersive 3D displays (5.1 vs. 5.9), $F(1, 21) = 4.57, p < .05$, and the more immersive 3D displays (5.1 vs. 6.0), $F(1, 21) = 4.81, p < .05$. However, there was no significant difference in review media usability ratings between the less immersive and more immersive 3D displays (5.9 vs. 6.0), $p > .05$.

3.2.2 Review Task Usability. Figure 7 also depicts the mean review task usability ratings for each experimental display. The ANOVA showed that display type affected review task usability ratings, $F(4, 35) = 14.08, p < .05$. Participants produced lower review task usability ratings with the 2D display than all 3D displays combined (5.0 vs. 5.9), $F(1, 35) = 4.69, p < .05$. Participants produced lower review task usability ratings with the 2D display than with the less immersive 3D displays (5.0 vs. 5.9), $F(1, 21) = 4.75, p < .05$, or the more immersive 3D displays (5.0 vs. 5.9), $F(1, 21) = 4.82, p < .05$. However, there was no significant difference in review task usability ratings between the less immersive and more immersive 3D displays (5.9 vs. 5.9), $p > .05$.

3.2.3 Quality of Collaborative Review Usability. Figure 7 shows the mean quality of collaborative review usability ratings for each experimental display.

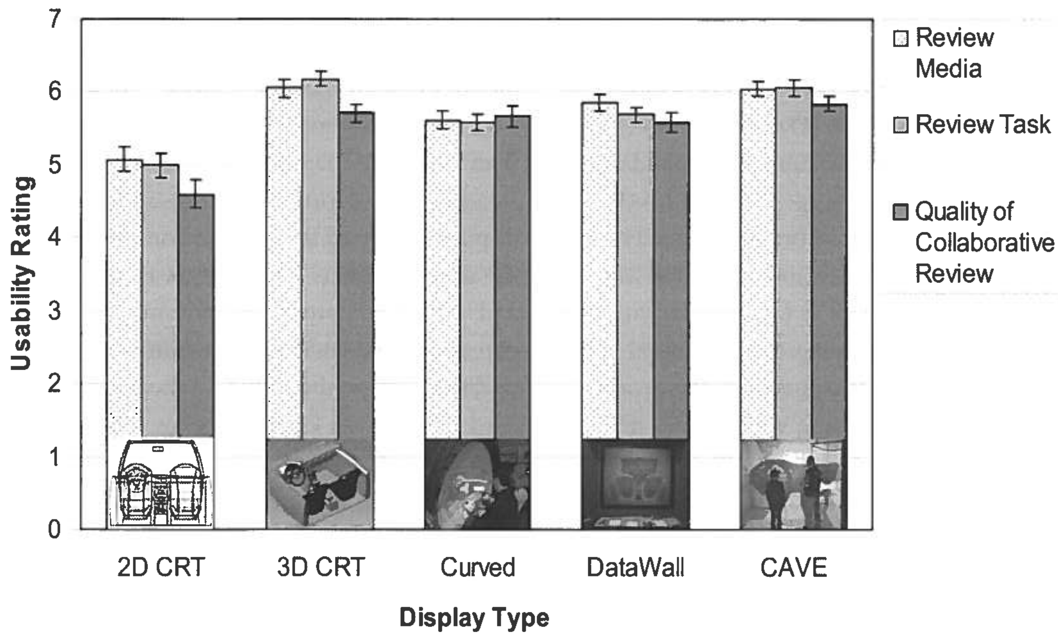


Figure 7. *Usability ratings for review media, review task, and quality of collaborative review as a function of display type.*

The ANOVA showed that display type affected quality of collaborative review usability quality ratings, $F(4, 35) = 11.74$, $p < .05$. Participants produced lower collaborative quality ratings with the 2D display than with all 3D displays combined (4.6 vs. 5.7), $F(1, 35) = 4.99$, $p < .05$. Participants produced lower collaborative quality ratings with the 2D display than with the less immersive 3D displays (4.6 vs. 5.7), $F(1, 21) = 5.21$, $p < .05$, or the more immersive 3D displays (4.6 vs. 5.7), $F(1, 21) = 5.13$, $p < .05$. However, there was no significant difference in collaborative review usability ratings between the less immersive and more immersive 3D displays (5.7 vs. 5.7), $p > .05$.

Informal observation indicated that participants had difficulty communicating with each other while wearing 3D glasses in the more immersive 3D displays (DataWall and CAVE). In the CAVE, most participants were observed having difficulty controlling the Pinch glove to manipulate the model, and reported this in their written comments. In some cases, they noted that the model was unintentionally manipulated when one participant viewed the model and the other accidentally moved it. Many CAVE participants had difficulty exploring the model through extensive interaction, and some of them used only one or two views of the model, perhaps assuming that the model was static. Some suggested that more training was needed for some CAVE components (i.e., 3D glasses, Pinch glove, and working within the four CAVE walls).

4 Discussion

We hypothesized that collaborative design review performance should be superior with 3D displays. Consistent with our hypothesis, participants detected more flaws with all 3D displays than the 2D CRT display. However, the time on task result contradicted the hypothesis because participants took more time with 3D CRT and Curved displays than the 2D CRT display. Thus, participants detected more flaws with the less immersive displays (3D CRT and Curved displays com-

bined) than the 2D display, but took more time, demonstrating a speed-accuracy trade-off. There was no difference in speed or accuracy between 2D displays and the more immersive 3D displays (DataWall and CAVE).

Participants may have had difficulty understanding and interpreting the 2D CAD model because it was static and provided little depth information (see Figures 2 and 3). The usability evaluation indicated that 2D participants rated the 2D model as less representative of real-world objects than those using the 3D model. The 2D participants rated it harder to perform the review task with the 2D model than those with the 3D displays. Participants in the 2D condition rated the facilitation of collaborations lower than those in 3D environments. Since participants could not explore the 2D model, they took less time and perhaps ran out of identifiable flaws, leading to fewer flaws detected.

One could argue that participants in the 2D CRT display were at a disadvantage because the CAD model could not be translated or rotated. However, sufficient resolution was provided so that all the design flaws could be detected in all displays. Thus, design flaws could be identified with the 2D model without translation or rotation. The lack of a translation or rotation function in the 2D CRT display should not have affected the comparison results.

The performance results did not support the hypothesis that performance should increase as the level of immersion increases for 3D displays. Among the 3D displays, participants detected fewer flaws (but spent less time) with the more immersive displays (DataWall and CAVE) than the less immersive 3D CRT and Curved displays. This is another speed-accuracy trade-off. Although the usability testing did not indicate differences between less immersive and more immersive 3D displays, informal observations and subjective reports indicated that the participants were not comfortable and confident performing the review task with the more immersive 3D displays (DataWall and CAVE). Dynamic manipulation of the model to change the viewpoint may have been difficult for inexperienced participants. DataWall and CAVE participants indicated that they did not have sufficient time to practice with novel components, such as the 3D eyewear and the Pinch glove (for the CAVE). CAVE participants reported that it was difficult to interactively control the model through fingertip manipulation with the Pinch glove. These factors might have led participants to terminate the trials more quickly than when the 3D CRT or the Curved display was used. In addition, wearing 3D glasses in DataWall and CAVE displays appeared to degrade communication between the team members. Difficulty with model manipulation and with team communication may have led to shorter detection time and lower usability ratings. A comparison of display costs (see Table 1) shows that using the standard CRT or the Curved display would probably be more cost-effective than using the DataWall or the CAVE display, as long as an increase in detection time did not present a serious obstacle.

5 Conclusion

This study compared the effectiveness of five display technologies for supporting a collaborative design review. The results indicated that neither the 2D display nor more immersive display technologies (DataWall and CAVE) were as useful as less immersive technologies (3D CRT and Curved displays) in terms of the number of design flaws detected. However, these advantages came at the cost of increased time to detect. The results may have been due to the lack of training for the participants with DataWall and CAVE displays. Usability testing results did not indicate an advantage for using more immersive display technologies over the less immersive technologies. The level of immersion in VEs should be appropriately scaled to optimize users' interaction with the display media and to facilitate collaboration.

It would appear that, with minimal training and experience, less immersive 3D displays may offer a cost-effective alternative for collaborative design review to either 2D displays or more immersive 3D display technologies. This advantage comes at an increased time cost for the collaborative design review. The relative importance of speed vs. accuracy depends on the application. In some cases, with strict deadlines, the time for a design review might be very short because every minute matters. We suspect that the cost of a few extra

minutes may be worthwhile in many cases given the importance of identifying design flaws that could have long-term implications for product success.

Acknowledgments

This research was conducted collaboratively by Defence R&D Canada—Toronto, CAE Professional Services—Canada, and the Virtual Environments Technologies Centre, Integrated Manufacturing Technologies Institute, of the National Research Council Canada, located in London, Ontario. The authors appreciate the support of the Synthetic Environments Coordination Office at the Canadian Forces Experimentation Centre (Ottawa). The authors are grateful for technical support from Mr. Niall Murray and Ms. Ellie Withers of the Integrated Manufacturing Technologies Institute. The authors also thank Mr. Franklin Lue and Mr. Leo Boutette of the Human Effectiveness Experimentation Centre of Defence R&D Canada—Toronto, for their technical support.

References

Barfield, W., Zeltzer, D., Sheridan, T., & Slater, M. (1995). Presence and performance within virtual environments. In W. Barfield & T. A. Furness III (Eds.), *Virtual environments and advanced interface design* (pp. 474–513). New York: Oxford University Press.

Bolstad, C. A., & Endsley, M. R. (1999). *Information dissonance, shared mental models and shared displays: An empirical evaluation of information dominance techniques* (AFRL-HE-WP-TR-1999-0213). Wright-Patterson AFB, OH: Air Force Research Laboratory.

Bowman, D. A., Datey, A., Young, S. R., Umer, F., & Omar, V. (2002). Empirical comparison of human behaviour and performance with different display devices for virtual environments. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 2134–2138). Santa Monica, CA: Human Factors and Ergonomics Society.

Brooks, F. (1999). What's real about virtual reality? *IEEE Computer Graphics & Applications, 19*(6), 16–27.

Commission of the European Communities. (2000). *Commission recommendation on safe and efficient in-vehicle information and communication systems: A European statement of principles on human machine interface.* Brussels, Belgium: European Union.

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '93* (pp. 135–142). New York: ACM.

Dede, C., Salzman, M. C., & Loftin, R. B. (1996). Maxwell World: Learning complex scientific concepts via immersion in virtual reality. *Proceedings of the 2nd International Conference on Learning Sciences* (pp. 22–29). New York: Academic Press.

Dynott, R., Lemoine, P., Piechulla W., Stefani, O., Wakula, J., & Williams, M. (2002). *Virtual reality systems for perceived ergonomic quality testing of driving task and design* (Tech. Rep. No. 3). Regensburg, Germany: Universität Regensburg, Traffic Research Group.

Green, P., Levison, W., Paelke, G., & Serafin, C. (1994). *Suggested human factors design guidelines for driver information systems* (Tech. Rep. UMTRI-93-21). Ann Arbor, MI: The University of Michigan, Transportation Research Institute.

Howell, D. C. (1999). *Fundamental statistics for the behavioral sciences* (4th ed). Pacific Grove, CA: Brooks/Cole.

Japan Automobile Manufacturers Association. (2000). *Guideline for in-vehicle display systems.* Tokyo, Japan: JAMA.

Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kocian, D. F., & Task, H. L. (1995). Visually coupled systems hardware and the human interface. In W. Barfield & T. A. Furness III (Eds.), *Virtual environments and advanced interface design* (pp. 175–257). New York: Oxford University Press.

McCormick, E. P., Wickens, C. D., Banks, R., & Yeh, M. (1998). Frame of reference effects on scientific visualization subtasks. *Human Factors, 40,* 443–451.

Olmos, O., Wickens, C. D., & Chudy, A. (2000). Tactical displays for combat awareness: An examination of dimensionality and frame of reference concepts and the application of cognitive engineering. *International Journal of Aviation Psychology, 10,* 247–271.

Owen, S. (1999). *HyperVIS: Teaching scientific visualisation using hypermedia.* Report to the Project of the ACM SIGGRAPH Education Committee, the National Science Foundation, and the Hypermedia and Visualization Laboratory, Georgia State University.

Peterson, B., Wells, M., Furness, T. A., III, & Hunt, E. (1998). The effects of the interface on navigation in virtual environments. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 1496–1500). Santa Monica, CA: Human Factors and Ergonomics Society.

Planchard, D. C., & Planchard, M. P. (2001). *Engineering design with SolidWorks 2001*. Mission, KS: Schroff Development Corporation.

Snow, M. P., & Williges, R. C. (1997). Empirical modeling of perceived presence in virtual environments using sequential exploratory techniques. In *Proceedings of the Human Factors and Ergonomic Society 41st Annual Meeting* (pp. 1224–1228). Santa Monica, CA: Human Factors and Ergonomics Society.

Tickoo, S. (2002). *Pro/Engineer for designers*. Schererville, IN: Cadcim Technologies.

Werkhoven, P. J., & Groen, J. (1998). Manipulation performance in interactive virtual environments. *Human Factors, 40*, 432–442.

Wickens, C. D. (2000). The when and how of using 2-D and 3-D displays for operational tasks. In *Proceedings of the IEA2000/HFES 2000 Congress, 3* (pp. 403–406). Santa Monica, CA: Human Factors and Ergonomics Society.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Thomas, L., Merlo, J., & Hah, S. (1999). Immersion and battlefield visualization: Does it influence cognitive tunnelling. In *Proceedings of the Third Annual Army Federated Lab Symposium*. College Park, MD: US Army Research Federated Lab Consortium.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21*, 229–244.